STA314: Statistical Methods for Machine Learning I (Fall 2018)

Instructor: Daniel Simpson

Office Hours: Monday 1-2pm **(Lecture room)**,
Tuesday **2:30-3:30pm**
Other times by appointment only

Email: simpson@utstat.toronto.edu

Email Policy: I'll try to get back to you within 2 work days, so if you email me at 7pm Wednesday, I'll get back by 5pm Friday. I don't check email on weekends.

Lectures:
Monday 11:00–13:00, BI 131 (Banting Institute)
Tuesday 13:00–14:00, AH 100 (Alumni Hall)

Tutorials (choose 1):
Friday 13:00–14:00 BA1230, BA1240, BA2165
Friday 12:00–13:00 SS1088, RW143

Assessment:

Final Exam: 55%

- Three Hours
- Covers entire course (1/3 from first half, 2/3 from second half)

Midsemester Exam: 25% **(12 November, 11:00am)**

- One Hour (in lecture time)
- No make ups. If you cannot make the exam, you get 0.
- If you would do better with only the final exam (as 80%), that will be your final mark.
- **The test will be written in a room other than the lecture room (location to be announced).**
- **If the test is missed for a valid reason, you must provide appropriate documentation, such as the University of Toronto Medical Certificate, University of Toronto Health Services Form, or College Registrar's Letter. You must submit this documentation within one week of the test. If documentation is not received in time, your test mark will be zero. If a test is missed for a valid reason, its weight will be shifted to the final exam.**

Homework: 20%

- 5 throughout the semester worth 4% each
- Material covered in these assignments will be assessable on the exams
- **Due 28 September, 5 October, 12 October, 26 October, 23 November. All assignments will be due at 12:00pm on the due day.**
- **There will be an optional 6th homework assignment due 12:00pm on 30 November. If you choose to do this, I will use the best five of the homework assignments when computing your final mark.**

Lateness policy:

Homeworks are due sharply at the appointed time and will receive significant penalties if late.

Re-grading policy:

Regrading requests should only be made for genuine grading errors, and should be initiated by writing or typing a complete explanation of your concern (together with your full name, student number, and e-mail address) on a separate piece of paper, and giving this together with your original unaltered homework/test paper to the instructor within one week of when the graded item was first available. Warning: your mark may end up going down rather than up.

Textbook and slides:

- Our text will be *An Introduction to Statistical Learning with Applications in R* by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani.
- The book is freely available online from this address: http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Sixth%20Printing.pdf
- Occasionally we will go beyond the textbook, in which case alternate references will be provided.
- Lecture slides will be available on quercus before each lecture.
- Tutorials will be available on quercus after the lecture on Tuesday.

Computing:

- The course will be run using the R computing environment.
- You are strongly encouraged to use RStudio (https://www.rstudio.com), which is a free IDE for R.
- All instructions in the course will assume that you have the latest version of both RStudio and R installed. We will not answer any R related questions unless both of these things are true.
- The best resource for R help is always google.